

Eur. J. Clin. Chem. Clin. Biochem.
Vol. 32, 1994, pp. 369–375
© 1994 Walter de Gruyter & Co.
Berlin · New York

Concentration-Dependent Profiles for Describing the Scatter of Results of Interlaboratory Surveys

By W. J. Geilenkeuser and G. Röhle

Institut für Klinische Biochemie der Universität Bonn, Germany

(Received August 24, 1992/February 23, 1994)

Summary: In an interlaboratory survey for the quantitative determination of a clinical chemical quantity, samples of the same specimen are analysed in different laboratories. If the number of participating laboratories is sufficiently large, then the differences between the 50th percentile (median) and e. g. the 25th and 75th percentiles of the results give a very reliable impression of the range of interlaboratory scatter for the particular analytical technique.

Results from a relatively large number of interlaboratory surveys, in which specimens containing different concentrations of the analyte are investigated, can be handled in the same way. If the resulting differences between the chosen percentiles are plotted against the median, and the corresponding two regression lines (upper and lower) are drawn, the results are asymmetric scatter profiles covering the concentration range of the specimen collective.

Numerous options are available. Thus, a profile's power of characterizing the scatter correctly can be improved by weighting of the results. Moreover, scatter profiles can be based on different variables of the survey, such as the analytical method, or the observation period, etc. They may be based on the total collective of all results for a given quantity, or they can be constructed for subcollectives of results obtained with a single analytical method. Further, it is possible to present the results of all subcollectives in a single pair of scatter profiles. This latter type of analysis provides profiles of the average scatter for a collective of different analytical methods, which are unaffected by any systematic differences that may exist between the methods.

1. Introduction

In addition to the accuracy of a quantitative analytical method, the reproducibility of replicated analyses, i. e. the precision, is also an important property of any analytical system.

The extent of scatter depends, inter alia, on the properties of the analysed material, on the concentration of the analyte and the circumstances under which analyses are repeated. Thus, a distinction is drawn between intra-run (with automated or mechanized analytical systems), intra-assay (within an analytical series), inter-assay (from series to series) and interlaboratory scatter for a given concentration of an analyte in a given matrix. Intra-run, intra-assay and inter-assay scatter are determined in a single laboratory, and they can be described with simple

parameters, provided all analyses are performed on the same specimen.

If an analyte is determined repeatedly in different specimens, a concentration-dependent increase or decrease in scatter (first described for clinical chemical analyses by Keller (1)) is often found. Precision profiles are well established (2, 3) for the documentation of the scatter of analytical results from different specimens containing different concentrations of a given analyte.

For the same analyte in the same specimen, factors affecting the variation of results from different laboratories, e. g. in an interlaboratory survey, are much more complex than those affecting variation within the same laboratory.

In addition to the scatter of values within each laboratory, the number and magnitude of random errors in an

interlaboratory survey also depend on the extent to which the handling procedures and the analytical apparatus can be standardized in different laboratories for the same analytical principle. In the first instance, this assumes the absence of analytical errors. Random errors of this kind arise only from the analytical principle; the inherent extent of these errors may vary for different analytical principles. The scatter of results from an interlaboratory survey may be influenced by other random factors that are independent of the analyst, e. g. differing matrices of the analytical samples, the composition of the collective of participants, or possibly by exceptional climatic conditions during the survey. The influence of these factors on an analytical result is to some extent systematic, but because of changing preconditions it is a matter of chance as how far it plays a role in a given survey. On the other hand, interlaboratory scatter is crucially affected by the systematic differences and errors of the individual laboratories. Such differences can arise when different laboratories use the same analytical principle but different analytical procedures; when different analytical principles are used, these differences are often very pronounced. These different causes of errors reveal different scatters from survey to survey and, in addition, the magnitude of these scatters is often dependent on the concentration of the analyte.

Under these complex circumstances it seemed useful to use graphs to illustrate coherencies between the scatter and some of the analytical prerequisites.

For this purpose, the present account describes the construction of concentration-dependent asymmetric scatter profiles from the data of interlaboratory surveys. Our approach is based on proposals for the construction of scatter profiles from survey results published in 1986 (4, 5) and now describes the construction procedure in more detail and with some useful extensions. It is our aim to describe the principle on which, for a given quantity, different scatter profiles can be calculated according to the question to be answered, and to illustrate these procedures with examples. The scatter profiles described here do not give any information on inaccuracy defined by target values. Owing to the well known fact that the interlaboratory scatter of results generally is much more affected by systematic errors than by random errors, it can only give a relative indication of the average differences of inaccuracy under defined conditions.

2. Materials and Methods

2.1 Interlaboratory surveys

The data for these investigations originate from interlaboratory surveys, which were conducted in recent years in collaboration with the German Society for Clinical Chemistry.

In an interlaboratory survey of quantitative analyses, each participating laboratory is sent N samples (often $N = 2$) of material (e. g. serum) for measurement of requested analytes. In a single interlaboratory survey, the returned results for each analyte form N collectives of analytical values. Where applicable, there will also be N collectives of results for each different analytical principle.

The distribution of these collectives is subject to the influencing factors described in the introduction.

2.1.1 Sample materials

Each sample consisted of a lyophilizate of pre-treated (e. g. elimination of lipids etc) serum, which had been supplemented with the appropriate analytes for use in the interlaboratory survey. These samples were prepared by commercial companies.

2.2 Data processing

All calculations and graphical evaluations were implemented in the programming language C running under a UNIX-system. The regression curves are calculated according to the well established least-squares method, see e. g. 1. c. (6–8).

2.3 Definitions and terms

The following symbols are used:

A represents a quantity;

M_A represents the set of methods applied to A in the interlaboratory survey;

$M \in M_A$ represents one method for A from the set M_A .

A *total collective* comprises all the analytical values for an analytical quantity A , obtained by analysing a given sample in an interlaboratory survey using all available methods M_A .

A *subcollective* comprises all the analytical values for an analytical quantity A , obtained by analysing a given sample in an interlaboratory survey using a single method $M \in M_A$.

$R = \{i_1, i_2, \dots, i_k\}$ represents the index set for the calculation of a scatter profile of selected interlaboratory surveys ($i_j, k \in \{1, 2, 3, \dots\}$).

P_n represents the n -th percentile of a collective of values ($n \in \{1, 2, \dots, 99\}$).

An *absolute scatter element*, E_m , of a collective of results is defined by the value triple (P_m, P_{50}, P_{100-m}) . ($m \in \{1, 2, \dots, 49\}$).

A *relative scatter element*, E_m^r , of a collective of results is defined by the value triple $(P_m^r, P_{50}^r, P_{100-m}^r)$, and it is derived from the absolute scatter element, E_m , by calculating the relative differences between the percentiles P_m and P_{100-m} and the median P_{50} :

$$E_m^r = ((P_m - P_{50})/P_{50}, P_{50}, (P_{100-m} - P_{50})/P_{50}).$$

$n_r, r \in R$ represents the number of analytical values in a collective of results in interlaboratory survey $r \in R$.

n_{\min} represents the minimal number of results required for the calculation of E_m , i. e. $n_r \geq n_{\min}$. This restriction is necessary, because the reliability of percentiles as quantities of scatter decreases in small collectives.

n_{\max} represents the largest value of n_r in the collective used for the construction of the scatter profile.

2.4 Calculation of the scatter profile

2.4.1 Basic model

The procedure for the calculation of a scatter profile for results obtained with a given method is as follows.

For the quantity A , the method M , and the number of selected inter-laboratory surveys R , all scatter elements E_m and the derived E'_m are calculated for $n_r \geq n_{\min}$.

Let N be the number of samples distributed in an interlaboratory survey, and K the number of interlaboratory surveys; the maximal number of scatter elements is then $N \cdot K$.

With all relative scatter elements, we get two sets of points, according to the relative differences between the chosen 'low' and 'high' percentiles and the corresponding medians. The resulting collectives contain a maximum of $N \cdot K$ points each, distributed above and below the abscissa. We denote these two sets of points with W^- and W^+ , respectively.

In order to characterize the "mean" scatter, it is necessary to calculate regression curves through both sets of points. The appropriate formula (e.g. linear, polynomial, hyperbolic) must be chosen to suit the particular circumstances.

The above procedure can be illustrated with figure 1a, which shows the scatter profile (hyperbolic regression) for the following analytical quantity and conditions:

quantity A	creatinine in serum;
method M	kinetic Jaffé method;
$R = \{901, 902, \dots, 928\}$	for 24 surveys (8 a year) between 1990 and 1992;
$m = 25$	i.e. the scatter elements consist of the percentiles P_{25} , P_{50} and P_{75} ;
$n_{\min} = 20$	i.e. collectives containing fewer than 20 results were not considered.

These values for m and n_{\min} apply to all subsequent examples of scatter profiles.

2.4.2 Variations and modifications of the basic model

2.4.2.1 Presentations combining different analytical principles

The basic model yields a scatter profile for the values of an analytical quantity, which are measured by using a single analytical principle. The results for an analytical quantity, A , obtained by using all available principles of analysis, $M \in M_A$, can be plotted in 2 ways.

2.4.2.1.1 Method-independent presentation

The scatter profile is calculated from the total collectives, irrespective of the analytical methods used. The calculation of the regression curves is therefore based on the $N \cdot K$ points (maximum) on either side of the abscissa, where K is the number of surveys and N is the number of samples for a survey.

2.4.2.1.2 Method-dependent presentation

The systematic differences between the individual methods are not taken into account in the scatter profile, in that the scatter elements for each subcollective are calculated separately. The final diagram is therefore a cumulative presentation of the scatter profiles for each individual $M \in M_A$. Depending on the number of analytical principles, $|M_A|$, the maximal number of points for calculation of the regression is therefore $|M_A| \cdot N \cdot K$.

Method-dependent and method-independent presentations can be compared by reference to figure 2 (a and b).

2.4.2.2 Weighting

The quantities, W^- and W^+ , can be weighted in accordance with the number of participants, as follows.

First, taking n_r as the number of participants, every $w \in W$ (for $W = W^- \cup W^+$) is assigned a weighting according to the formula: $g_r = (n_r/n_{\max}) \cdot 40$, where g_r is the weighting factor. The point w is then incorporated g_r -times into W to give the set W' , which then serves as the basis for calculation of the final regression curve. (To realize up to 40 different levels of weighting may seem somewhat arbitrary, but this procedure represents a good compromise between the computational effort and the resulting regression curves, because there would be no relevant change in the regression formulae if we took the actual number of participants into account.) In the diagrams weighting is roughly indicated by different sized dots. The regression follows any resulting shift of the centres. An example is given in figure 3 (a and b).

3. Discussion

3.1 Prerequisites for the scatter profiles

It is expected that the evaluation of an interlaboratory survey provides information about the actual state of the analytical art for a given quantity. In this context, the range in the scatter of the results from different laboratories is one of the essential aspects.

To describe the asymmetric scatter of a given collective of survey results, we chose as a basic measure the differences between the median and the 25th percentile and the 75th percentile, respectively. It was decisive for this choice that, one, non-normal distributed values can be characterized more reliably by the aid of percentiles than by the often used standard deviation and, two, limiting the values included to the 25th and 75th percentiles is a better basis for evaluating the state of the analytical art than using all the values of a collective, since the values outside the range chosen can be assumed to be more often affected by disproportionate systematic errors.

In spite of this limitation to half of the collectives, the scatter of nearly equal collectives can differ from survey to survey even if the same specimen is analysed. Furthermore, the known fact that the analytical scatter can change with the concentration of an analyte must be taken into account. Then the simple solution would be to plot the scatters of survey results observed within a defined space of time versus the corresponding medians, and to calculate the appropriate regression curves. Such profiles show the average interlaboratory scatter for a larger concentration range. Thus they give more information on the state of the analytical art than do the separate evaluations of interlaboratory surveys.

3.2 Basic model

Scatter profiles, constructed according to the procedure described in section 2.4.1, give a clear graphical display

of the extent, the asymmetry and the concentration-dependency of the interlaboratory scatter for a given combination of analyte and analytical principle or method. They enable an easy comparison of different analytical principles with respect to the above-mentioned characteristics. Thus the examples of figure 1 show that the interlaboratory scatter of the results for creatinine determined with the kinetic method of Merck (fig. 1a) is significantly larger than that determined with the enzymatic method of Boehringer-Mannheim (fig. 1b, tab. 1).

Criteria for justifying the use of a particular regression method are difficult to state generally. In most cases,

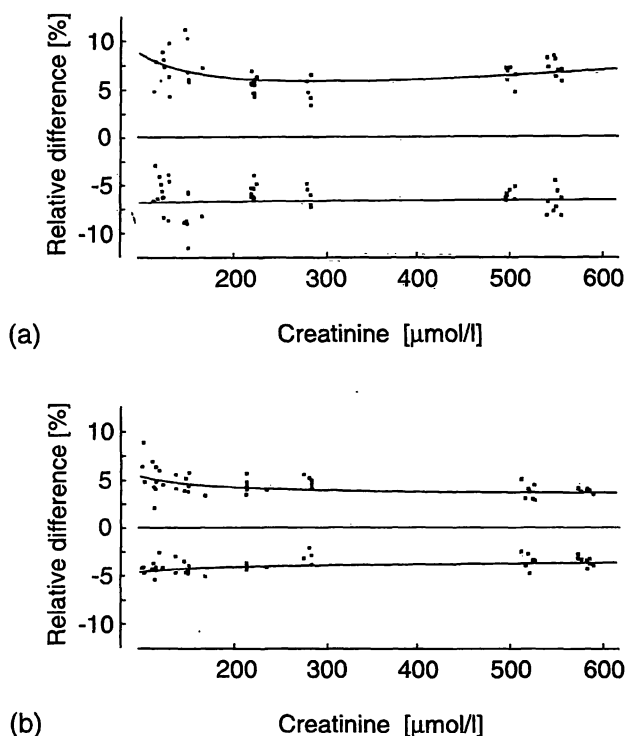


Fig. 1 Scatter profiles (hyperbolic regression, calculated from the 25th and 75th percentiles, respectively) of the results of creatinine determinations in 24 interlaboratory surveys (1990–1992) performed with

a) the kinetic *Jaffé* method of Merck, Darmstadt (on a basis of 3100 analytical results)

b) the enzymatic method of Boehringer-Mannheim (on a basis of 3700 analytical results)

however, the regression curve with the best mathematical approximation will also be suitable from the analytical point of view. The hyperbolic regression, however, appears to be the most suitable, due to its ability to reproduce the sometimes marked increase of scatter for small analytical values.

3.3 Presentations combining different analytical principles

A method-independent scatter profile (see section 2.4.2.1.1) includes the systematic differences and errors of the different analytical principles. This type of profile (fig. 2a) compared with the profiles of figure 1 or those of figures 2b or 2c (see tab. 2) therefore gives a visual impression of whether the values obtained by different analytical principles are homogeneous, or whether the interlaboratory scatter is increased by serious systematic differences. In the latter case, the total scatter is larger than that for most of the separate analytical principles.

A method-dependent scatter profile (see section 2.4.2.1.2) does not take account of these systematic differences between analytical principles, and it rather represents a measure of the “average” scatter of all the analytical principles used in the survey (fig. 2b). For quantities, whose analytical values are subject to only small systematic differences, there is little difference between the method-dependent and method-independent scatter profiles.

Figures 2a and 2b, show these two types of scatter profile for the determination of creatinine in serum in the period 1990–1992. In the method-independent profile (fig. 2a), systematic differences between different analytical principles give rise to a relatively wide separation of the regression lines from the abscissa, especially in the lower concentration range.

The calculation of the method-dependent profiles according to section 2.4.2.1.2 (fig. 2b) treats small subcollective of results determined with the same method with the same weight as large ones. How frequently distinct

Tab. 1 Differences in % between abscissa and the two regression curves, lower and upper, for the 25th percentiles and the 75th percentiles, respectively.

Creatinine [µmol/l]	Lower regression curve					Upper regression curve				
	figures					figures				
	1a	1b	2a	2b	2c	1a	1b	2a	2b	2c
100	-6.8	-4.6	-9.3	-6.8	-7.0	8.6	5.3	8.5	7.9	7.4
200	-6.7	-4.1	-7.6	-5.9	-5.7	6.2	4.2	6.5	5.5	5.2
300	-6.7	-3.9	-6.8	-5.6	-5.3	5.8	3.9	5.7	4.9	4.6
400	-6.7	-3.9	-6.3	-5.5	-5.2	5.9	3.7	5.2	4.7	4.5
500	-6.7	-3.8	-5.9	-5.5	-5.1	6.3	3.6	4.8	4.7	4.5

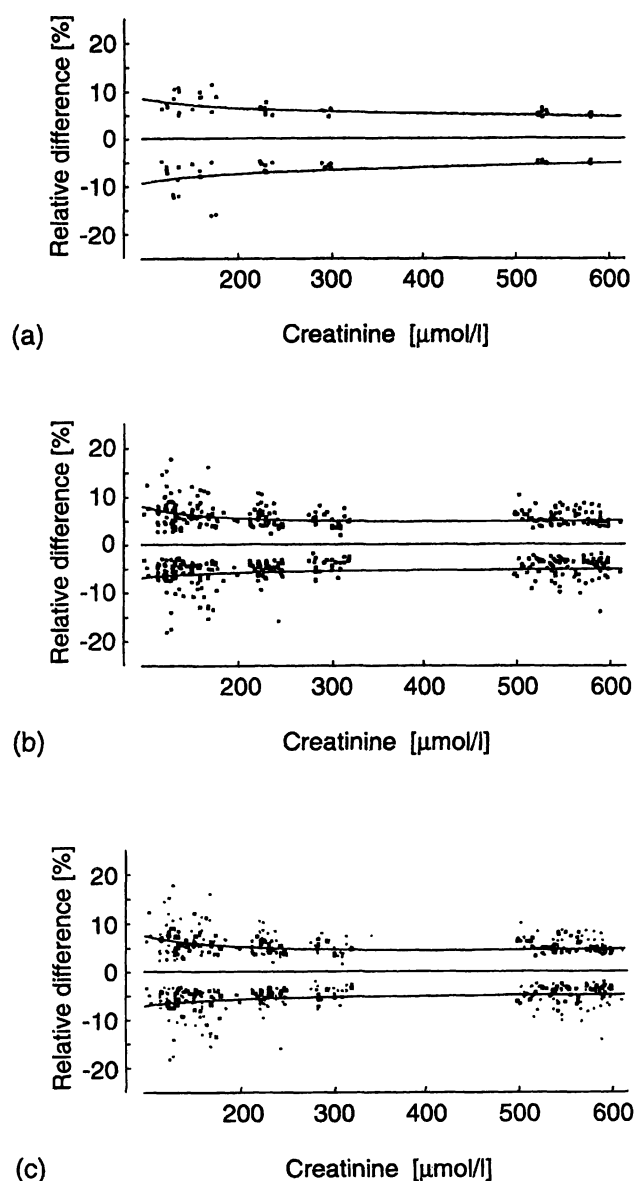


Fig. 2 Scatter profiles (hyperbolic regression, calculated from the 25th and 75th percentiles, respectively) of the results of creatinine determinations (including all methods used) in 24 interlaboratory surveys conducted between 1990 and 1992 on a basis of 25 100 analytical results.

- a) method-independent;
b) method-dependent;
c) method-dependent and weighted.

methods are used determines in part the state of the analytical art. In this case it is reasonable to weight the scatter elements according to the size of the subcollectives.

3.4 Weighting

The number of participants measuring a particular analytical quantity may differ greatly from one interlaboratory survey to another. Moreover, if the values obtained with different analytical principles are included in a single method-dependent presentation (section 2.4.2.1.2), the number of participants in each collective

can vary. The reliability of the estimation of scatter by means of percentiles increases, however, with the number of values available for the calculation.

In the construction of scatter profiles, each plotted point should have the same statistical status, which is not the case if the underlying collectives of results are of unequal sizes. This problem is overcome by a calculation of weighting, as described in section 2.4.2.2, which gives a higher weighting to collectives containing larger numbers of results.

As an example, figure 2c shows the scatter profile of figure 2b after weighting. Depending on the number of different methods, the number of points in a method-dependent scatter profile can be much higher compared with the method-independent presentation. In this particular case, the analytical values are similar for large and small collectives, so that weighting has only a slight effect. In contrast, the method-dependent profile in figure 3a and 3b (measurement of potassium in serum in the period 1990 to 1992) shows the need for weighting. The total scatter (fig. 3a) is increased by the results of little-used, relatively inaccurate analytical methods. After weighting (fig. 3b), their influence is reduced and total scatter is about a fifth less, as shown clearly by table 2.

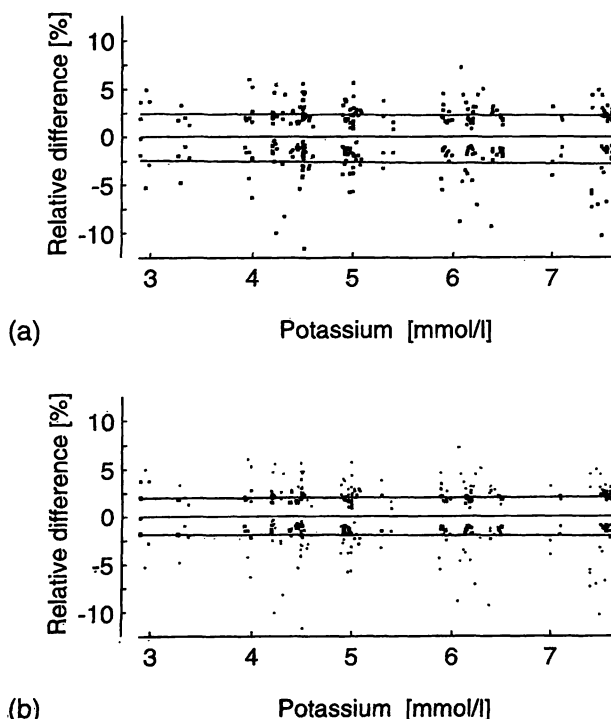


Fig. 3 Method-dependent scatter profiles (linear regression, calculated from the 25th and 75th percentiles, respectively) of the results of potassium determinations (including all methods used) in 24 interlaboratory surveys conducted between 1990 and 1992 on a basis of 17 100 analytical results.

- a) unweighted;
b) weighted.

Tab. 2 Differences in % between abscissa and the two regression curves, lower and upper, for the 25th percentiles and the 75th percentiles, respectively.

Potassium [mmol/l]	Lower regression curve		Upper regression curve	
	figures		figures	
	3a	3b	3a	3b
3	-2.4	-1.9	2.4	1.9
4	-2.5	-2.0	2.4	1.9
5	-2.6	-2.0	2.3	1.9
6	-2.7	-2.1	2.2	1.9
7	-2.8	-2.2	2.1	1.8

3.5 Choice of the appropriate scatter profile

There are many different possible scatter profiles for a given analytical quantity, depending on the options chosen for the calculation, namely

- observation period,
- analytical method or reagent combination (kit),
- alternative presentations with respect to methodology,
 - a) method-independent
 - b) method-dependent
- weighting,
- type of regression curve, and
- minimal size of collective.

It is generally useful to apply weighting. If it happens that the corresponding potential interferences are not present, the results are unaffected by the application of weighting. The choice of other options depends on the problem in hand and the information desired.

Though the scatter profiles of the type described here demonstrate primarily a limited aspect of the state of the analytical art, the information obtainable (from these profiles) was used to formulate analytical goals for the internal and external quality assessment:

The provisional regulations of the German Bundesärztekammer (9), under which the limits of inaccuracy and imprecision for the determination of 6 hormones were redefined, are based on scatter profiles calculated according to the following conditions ("parameters"):

- observation period 1987–1990,
- method-independent presentation,
- with weighting,
- hyperbolic regression with respect to the 25th and 75th percentiles.

As an example, the scatter profiles of progesterone are shown in figure 4a and table 3, respectively. The official limits published in 1992 (9) were the result of, on the

one hand, increasing the distance of the regression lines from the abscissa by a factor of 3, and, on the other hand, of fairly substantial simplifications.

The following considerations were important in the choice of these conditions, in particular the choice of the method-independent presentation. According to the basic guidelines of the German Bundesärztekammer of 1988 (10), the maximal acceptable inaccuracy for the 6 hormones is given by the reference method value and its assigned tolerance range, irrespective of the routine method used. Since this tolerance range is equally valid

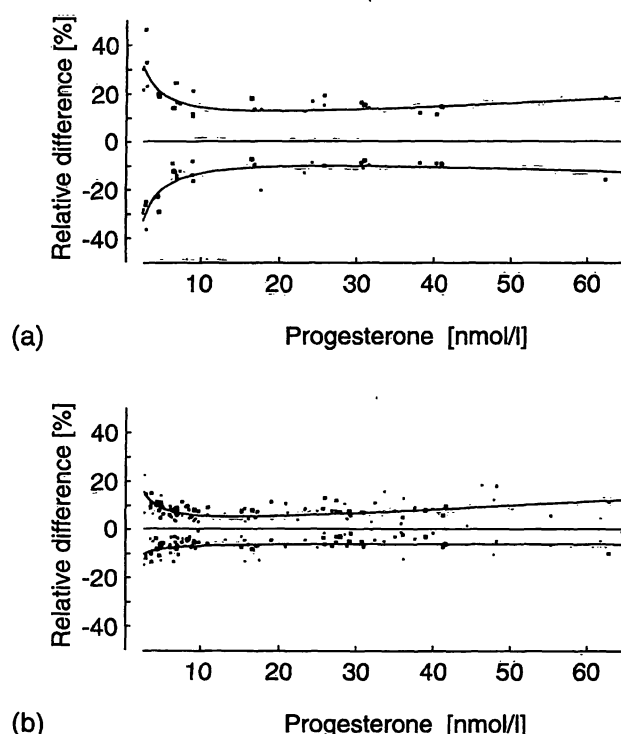


Fig. 4 Scatter profiles (hyperbolic regression, calculated from the 25th and 75th percentiles, respectively) of the results of progesterone determinations (including all methods used) in 16 interlaboratory surveys conducted between 1988 and 1991 on a basis of 2900 analytical results.

- a) method-independent;
- b) method(= kit)-dependent.

Tab. 3 Differences in % between abscissa and the two regression curves, lower and upper, for the 25th percentiles and the 75th percentiles, respectively.

Progesterone [nmol/l]	Lower regression curve		Upper regression curve	
	figures		figures	
	4a	4b	4a	4b
5	-19.9	-7.7	20.1	8.8
10	-13.1	-6.7	14.4	5.9
20	-10.6	-6.3	12.7	5.7
30	-10.5	-6.3	13.2	6.7
40	-11.1	-6.5	14.3	8.0

for all results, irrespective of the routine method used, it would be logical for it not to equalize the method-dependent systematic differences of the results as is the case for method (kit)-dependent profiles (see below).

In contrast, the method-dependent profile was chosen to establish the maximal acceptable imprecision (9) with otherwise no change in the "parameterization" of the scatter profile. Figure 4b shows the resulting profiles for progesterone. (The official published limits represent in principle approximately half the distance between the 2

profiles.) There is no direct relationship with internal laboratory precision. In view of the lack of generally accepted data for the internal laboratory control of precision, this procedure, however, represents a plausible alternative solution. The chosen criterion is therefore de facto the average scatter of the better half of the results from users of the same kit. Since this quantity includes the systematic differences between the different laboratories using the same kit, the resulting maximal acceptable imprecision represents a rather generous criterion for internal laboratory precision.

4. References

1. Keller, H. (1968) Über die konzentrationsbedingte Abhängigkeit von Streuung und Richtigkeit bei klinisch-chemischen Bestimmungen. In: *Automatisierung des klinischen Laboratoriums* (Grieser, G. & Wagner, G., eds.) Schattauer-Verlag Stuttgart.
2. Ekins, R. P. (1983) The precision profile: Its use in assay design, assessment and quality control. In: *Immunoassays for Clinical Chemistry*, 2nd ed. (Hunter, W. M. & Corrie, J. E. T., eds.) Churchill Livingstone, Edinburgh, pp. 76–105.
3. Ekins, R. P. (1976) General principles of hormone assay. In: *Hormone Assays and Their Clinical Application* (Loraine, I. A. & Bell, E. T., eds.) Churchill Livingstone, Edinburgh, pp. 2–72.
4. Röhle, G. & Voigt, U. (1986) Concept for the evaluation of analytical results in clinical chemistry. Part I: Concentration-dependent dispersion profiles of survey results. *Commun. Lab. Med.* 2, 81–85.
5. Röhle, G., Voigt, U. & Siekmann, L. (1986) Concept for the evaluation of analytical results in clinical chemistry. Part II: The accumulated distribution-free dispersion parameter and its application. *Commun. Lab. Med.* 2, 131–138.
6. Sachs, L. (1974) *Angewandte Statistik*. Springer-Verlag, Berlin Heidelberg New York, pp. 344–351.
7. Bortz, J. (1989) *Statistik für Sozialwissenschaftler*. Springer-Verlag, Berlin Heidelberg New York, pp. 215–248.
8. Armitage, P. (1973) *Statistical Methods in Medical Research*. Blackwell Scientific Publications, Oxford London Edinburgh Melbourne, pp. 150–156 and pp. 324–330.
9. Übergangsregelungen für die "Richtlinien der Bundesärztekammer zur Qualitätssicherung in medizinischen Laboratorien" (1992) *Dt. Ärztebl.* 89, B337–B338.
10. Qualitätssicherung der quantitativen Bestimmungen im Laboratorium (1988) *Dt. Ärztebl.* 85, A697–A712.

Dr. Wolf J. Geilenkeuser
Institut für Klinische Biochemie
der Universität Bonn
Sigmund-Freud-Straße 25
D-53105 Bonn
Germany

